The Battle of the Neighborhoods (2)

Segmenting and Clustering the city of Edinburgh

Business Problem

- Chain of bakeries planning to expand to Scotland
- Two large Scottish cities:
  - Edinburgh selected – festival foot traffic
  - Favourable rates & policies
  - Potentially cheaper location

Data

- '2020-2 Scottish Postcode Directory Files' dataset
- Provided by National Records Scotland
- Location of interest: Edinburgh

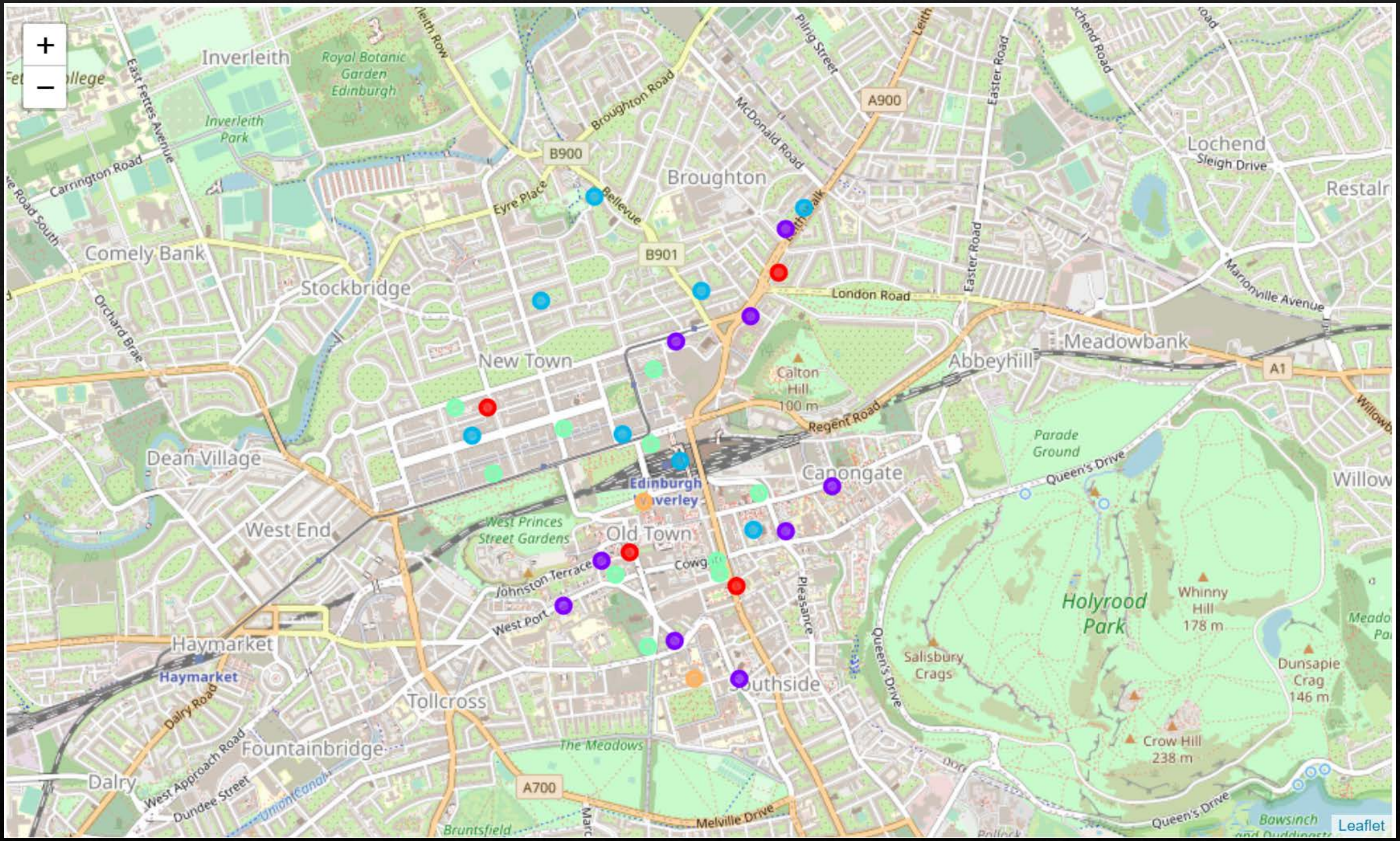
- Foursquare API

Methods

- Python packages
  - Pandas – using dataframes
  - NumPy – using arrays
  - Sci-kit learn – ML algorithm
  - Geopy – obtain coordinates
  - Folium – create maps
- K-Means clustering algorithm
- Foursquare API

Results

Cluster Label	Key types of Shop/Venue	Colour Legend
0.0	Gourmet Shop	Red
1.0	Ice Cream Shop; Tea Shop	Purple
2.0	Bakery	Blue
3.0	Creperie; Dessert Shop; Tea Room	Green
4.0	Gelato Shop	Orange


Inverleith
Royal Botanic Garden Edinburgh
Inverleith Park
Broughton
Lochend
Comely Bank
Stockbridge
New Town
Meadowbank
Abbeyhill
Calton Hill 100 m
Dean Village
Canongate
Edinburgh Waverley
West End
West Princes Street Gardens
Old Town
Parade Ground
Holyrood Park
Whinny Hill 178 m
Salisbury Crags
Dunsapie Crag 146 m
Crow Hill 238 m
Haymarket
Tollcross
Fountainbridge
Dalry
Southside
The Meadows
Bruntsfield
Broughton Road
Bellevue
Eyre Place
McDonald Road
Pilrig Street
Easter Road
Leith Walk
Carrington Road
East Fettes Avenue
Orchard Brae
London Road
Regent Road
Queen's Drive
Johnston Terrace
West Port
Cowgate
Pleasance
Dalry Road
West Approach Road
Dundee Street
Melville Drive
Marionville Avenue
Sleigh Drive
A900
B900
B901
A1
A700
Leaflet

Challenges faced...

- Number of clusters were not optimised
  - K-Means clustering
- Size of the original dataset
  - Limited number of Foursquare calls
- Coding inconsistencies/errors in Foursquare venues
  - Thai restaurant?

Recommendations

- The best location for a new bakery would be away from the blue and green clusters on the map!